На правах рукописи

КРИБЕЛЬ АЛЕКСАНДР МИХАЙЛОВИЧ

ВЫЯВЛЕНИЕ АНОМАЛИЙ И КЛАССИФИКАЦИЯ КОМПЬЮТЕРНЫХ АТАК В СЕТИ ПЕРЕДАЧИ ДАННЫХ НА ОСНОВЕ ПРИМЕНЕНИЯ ФРАКТАЛЬНОГО АНАЛИЗА И МЕТОДОВ МАШИННОГО ОБУЧЕНИЯ

Специальность 2.3.6
Методы и системы защиты информации, информационная безопасность

А В Т О Р Е Ф Е Р А Т
диссертации на соискание ученой степени кандидата технических наук

Санкт-Петербург – 2023

Работа выполнена в Федеральном государственном бюджетном учреждении науки «Санкт-Петербургский Федеральный исследовательский центр Российской академии наук» (СПб ФИЦ РАН) в лаборатории проблем компьютерной безопасности.

Научный руководитель	ЛАУТА Олег Сергеевич, доктор технических наук, профессор, профессор кафедры «Комплексного обеспечения информационной безопасности», ФГБОУ ВО «Государственный университет морского и речного флота имени адмирала С.О. Макарова».
Официальные оппоненты:	ГРЕЧИШНИКОВ Евгений Владимирович, доктор технических наук, профессор, директор по инновационному развитию, АО «Научно-исследовательский институт «Рубин».
	БИРЮКОВ Денис Николаевич, доктор технических наук, профессор, начальник кафедры, ФГБВОУ ВО «Военно-космическая академия имени А.Ф. Можайского».
Ведущая организация	Акционерное общество «Информационные технологии и коммуникационные системы», 127273, г. Москва, ул. Отрадная, 2Б стр. 1.

Защита состоится «18» мая 2023 г. в 14 часов 00 минут на заседании диссертационного совета 24.1.206.01, созданного на базе Федерального государственного бюджетного учреждения науки «Санкт-Петербургский Федеральный исследовательский центр Российской академии наук» (СПб ФИЦ РАН) по адресу: 199178, Санкт-Петербург, 14-я линия В.О., 39, каб. 401, e-mail dc@spcras.ru. Факс: (812)-328-44-50, тел: (812)-328-34-11.

С диссертацией можно ознакомиться в отделе аспирантуры (каб. 402а) Федерального государственного бюджетного учреждения науки «Санкт-Петербургский Федеральный исследовательский центр Российской академии наук» и на сайте http://www.spiiras.nw.ru/dissovet

Автореферат разослан «12» апреля 2023 г.

Ученый секретарь
диссертационного совета 24.1.206.01,
кандидат технических наук — Абрамов М.В.

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность темы. Темпы, с которыми развивается современная сфера информационных технологий, подвергает мировое сообщество целому ряду беспрецедентных угроз и уязвимостей, которые злоумышленнику открывают возможности к реализации компьютерных атак (КА).

Компьютерные атаки представляют собой сложное комплексное воздействие на сеть, в результате которого осуществляется их компрометации и нарушается управление процессами в сети передачи данных. Зачастую этому предшествует долгая и кропотливая работа: компьютерная техническая разведка, поиск характерных уязвимостей и захват информационных активов. Воздействие компьютерных атак возможно за счет использования технологий сбора информации, малоэффективных механизмов защиты, эксплуатации устаревших сетевых служб, протоколов и операционных систем.

Использование подобных уязвимостей даёт злоумышленнику возможность несанкционированно авторизовываться в системе, повышать пользовательские привилегии, изменять настройки сетевых устройств, прослушивать и перенаправлять трафик, блокировать сетевое взаимодействие и нарушать информационный обмен в сетях передачи данных (СПД).

Воздействия компьютерных атак приводит к появлению в сетях передачи данных аномальной активности трафика. Для постоянного мониторинга и обнаружения аномальной активности трафика в сетях передачи данных необходимо учитывать наличие большого количества сетевых маршрутов, на которых периодически возникают резкие колебания, задержки в передаче данных и большие потери пакетов, появление новых свойств сетевого трафика, а также необходимость обеспечения высокого качества обслуживания приложений. Все это послужило поводом для поиска новых методов обнаружения и прогнозирования компьютерных атак, к числу которых можно отнести машинное обучение и фрактальный анализ.

Важность и значимость решаемой задачи заключается в том, что на основе экспериментальных исследований возможно обосновать наилучший метод определения самоподобия для нестационарных и стационарных временных рядов, позволяющий с высокой точностью и достаточно быстро обнаруживать изменения в сетевом трафике сетей передачи данных, а также определить 7 структуру сети LSTM, позволяющую с высокой точностью и достаточно быстро прогнозировать факт воздействия компьютерных атак, на основе которого в дальнейшем могут вырабатываться проактивные мероприятия защиты.

Степень разработанности темы. В настоящее время вопросы, связанные с изучением самоподобных свойств временных рядов и их практиче-

ским применением в различных системах мониторинга, находятся в фокусе внимания многих исследователей. Фрактальные свойства исследуются во многих работах. Так, в работе Raimundo M.S. метод R/S-анализа используется для выявления закономерностей во временных рядах. В работе Dang T.D. моделируется VoIP-трафик, а также исследуются его фрактальные свойства. В работе Sánchez-Granero M.J. изучался не только показатель Херста, но и фрактальная размерность.

В тоже время большое внимание вопросам противодействия КА, выявлению аномалий и классификации КА в СПД уделяется такими исследователями как Д.А. Губанов, И.В. Котенко, М.В. Литвиненко, Д.А. Новиков, И.Б. Саенко, А.Л. Тулупьев, Д.Ю. Турдаков, А.А. Чечулин, А.Г. Чхартишвили, A.L. Barabasi, X. Zheng и др.

Учитывая требования нормативно-правовых документов, а также несмотря на сделанный учеными существенный задел, проблема выявления аномалий, раннего обнаружения известных и неизвестных компьютерных атак в сетях передачи данных, их классификации не может считаться разрешенной и требует проведения новых исследований

Таким образом, сложилось **противоречие** между возросшими деструктивными возможностями новых видов КА на СПД, приводящих к аномальной активности трафика, и устаревшими подходами к их выявлению в СПД, которое и предопределило выбор объекта и предмета исследования.

Цель исследования. Целью диссертационной работы является повышение эффективности выявления аномалий и классификации компьютерных атак в сетевом трафике сети передачи данных.

Научная задача заключается в разработке аналитической модели и методик выявления аномалий и классификация компьютерных атак в сетевом трафике сети передачи данных на основе применения фрактального анализа и методов машинного обучения.

Для достижения данной цели в диссертационной работе поставлены и решены следующие **частные задачи**:

1) анализ существующих моделей воздействия компьютерных атак в сети передачи данных;

2) анализ существующих алгоритмов выявления компьютерных атак, систем мониторинга и методик противодействия компьютерным атакам в сети передачи данных;

3) разработка аналитической модели выявления аномалий в сетевом трафике сети передачи данных в условиях компьютерных атак;

4) разработка методики раннего обнаружения аномалий в сетевом трафике сети передачи данных;

5) разработка методики классификации компьютерных атак в сетевом трафике сети передачи данных;

6) разработка архитектуры и программных прототипов компонентов системы раннего обнаружения и классификации компьютерных атак в сетевом трафике сети передачи данных, экспериментальная и теоретическая оценка эффективности предложенных моделей, методик и архитектуры.

Объектом исследования является СПД в условиях КА, а **предметом исследования** – модели, методики и алгоритмы выявления аномалий и классификации КА в СПД.

Научная новизна результатов исследования заключается в том, что:

разработанная аналитическая модель выявления аномалий в сетевом трафике СПД в условиях КА, в отличие от известных, описывает не только стационарный, но и нестационарный сетевой трафик и обосновывает выбор фрактального метода выявления аномалий в интересах дальнейшей классификации компьютерных атак в зависимости от типа трафика;

разработанная методика раннего обнаружения аномалий в сетевом трафике СПД, в отличие от известных, способна проводить прогнозирование и обнаружение не только известных, но и неизвестных КА на раннем этапе их проявления благодаря совместному применению методов фрактального анализа и искусственной нейронной сети LSTM-типа (рекуррентные нейронные сети с долгой краткосрочной памятью), в результате чего существенно сокращается время выявления аномалий и уменьшается количество ложных срабатываний;

разработанная методика классификации КА в сетевом трафике СПД отличается от известных тем, что в ней обнаружение компьютерных атак производится с использованием генеративно-состязательной сети, которая производит дообучение классификатора на скрытых латентных представлениях, полученных автокодировщиком, исходя из анализа данных о функционировании СПД;

предложенная архитектура и программные прототипы компонентов системы раннего обнаружения и классификации КА в СПД отличаются от известных тем, что они ориентированы на раннее обнаружение как известных, так и неизвестных КА с минимальным количеством ложных срабатываний за счет реализации методов фрактального анализа и искусственной нейронной сети LSTM-типа.

Теоретическая и практическая значимость результатов исследования заключается в том, что разработанные аналитическая модель и методики представляют собой научно-методическую основу, практическая реализация которой позволяет описать различные типы трафика в СПД, определять аномальные активности, основываясь на принципах самоподобия, и, исходя из типа трафика с применением различных методов машинного обучения, выявлять КА. Разработанные методики являются математической основой системы раннего обнаружения КА, основанные на

обнаружении аномалий в СПД и принятии эффективных мероприятий по защите с применением нейронной сети автокодировщика, состоящей из ячеек с долгой краткосрочной памятью и управляющим рекуррентным блоком. При этом повышается оперативность, точность и полнота выявления аномалий в СПД, что позволяет на практике эффективно применять разработанный подход в системах глубокой проверки сетевых пакетов в СПД.

Методология и методы исследования. В качестве математических положений исследования использованы: фрактальный анализ; методы машинного обучения; теория и практика систем связи; теория и практика проведения тестирования на проникновение; аналитико-статистические методы.

Положения, выносимые на защиту:

1. Аналитическая модель выявления аномалий в сетевом трафике СПД в условиях КА.
2. Методика раннего обнаружения аномалий в сетевом трафике СПД.
3. Методика классификации КА в сетевом трафике СПД.
4. Архитектура и программные компоненты системы раннего обнаружения и классификации КА в сетевом трафике СПД.

Обоснованность и достоверность результатов диссертационной работы подтверждается результатами вычислительных экспериментов, их сравнением с результатами других исследователей, практической апробацией разработанных модели и методик, а также одобрением основных положений диссертации на научно-технических конференциях, публикациями в ведущих рецензируемых журналах, внедрением результатов работы.

Апробация результатов. Основные научные и практические результаты работы докладывались и обсуждались на 10 научно-технических и научно-практических конференциях с 2017 по 2022 гг., к основным из которых относятся: Международная научно-практическая конференция «РусКрипто» (Московская область, 2021 и 2022); Межвузовская научно-практическая конференция «Актуальные проблемы обеспечения информационной безопасности» (Самара, 2017); Двенадцатая общероссийская молодежная научно-техническая конференция «Молодежь. Техника. Космос.» (Санкт-Петербург, 2020); Всероссийская научно-техническая конференция «Состояние и перспективы развития современной науки по направлению «Информационная безопасность»» (Анапа, 2021 и 2020); Всероссийская научно-техническая конференция «Состояние и перспективы развития современной науки по направлению «АСУ, информационно-телекоммуникационные системы»» (Анапа, 2021); Всероссийская научно-практическая конференция РАРАН «Актуальные проблемы защиты и безопасности» (Санкт-Петербург, 2019); Международная научно-

техническая и научно-методическая конференция «Актуальные проблемы инфотелекоммуникаций в науке и образовании» (Санкт-Петербург, 2019); Всеармейская научно-практическая конференция «Инновационная деятельность в Вооруженных Силах Российской Федерации», (Санкт-Петербург, 2017).

Внедрение и реализация результатов исследования. Результаты проведенного исследования нашли практическое применение в разработках, в которых автор принимал личное участие. О реализации основных результатов проведенного исследования имеются 3 акта о реализации в НИР «Корвет», «Потенциал-2018», «Свертка-СМ» (ФГКВОУ ВО «Военная орденов Жукова и Ленина Краснознаменная академия связи им. С.М. Буденного») и 1 акт о реализации в ОКР «Опорник» (ПАО «Информационные и телекоммуникационные технологии».

Публикации по теме исследования. По тематике диссертации опубликованы 17 работ, среди которых шесть статей, индексируемые в международных базах данных Web of Science и/или Scopus; десять статей в рецензируемых научных изданиях, входящих в перечень, установленный Министерством науки и высшего образования Российской Федерации; одно свидетельство о государственной регистрации программы для ЭВМ.

Структура и объем диссертации. Диссертационная работа состоит из введения, четырех разделов, заключения, списка сокращений, списка литературы. Работа выполнена на 145 страницах, выполненных печатным способом, содержит 77 рисунка, 5 таблиц и 4 приложения.

ОСНОВНОЕ СОДЕРЖАНИЕ РАБОТЫ

Во введении обоснована актуальность темы и определена цель диссертационной работы. Сформулированы основные научные положения исследования, выносимые на защиту, представлены сведения об апробации и публикациях по теме исследования, приведена аннотация диссертации по разделам.

В первом разделе проведен анализ устойчивости СПД и рассмотрены требования, предъявляемые к ее качеству. Проведен анализ информационных воздействий злоумышленника как одного из основных факторов, влияющих на состояние СПД. Рассмотрены существующие средства и способы защиты СПД. Проведен анализ существующих научно-методических положений об оценке устойчивости СПД, а также показаны их достоинства и недостатки, сформулированы задачи исследования. Сформулирована научная задача.

Второй раздел посвящен описанию разработки аналитической модели выявления аномалий в сетевом трафике СПД в условиях КА (рис. 1), предназначенная для описания сетевого трафика сразу двух типов: стационарного и нестационарного.

Для проверки стационарности ряда используется обобщенный тест Дики-Фуллера. Для определения аномальной активности в СПД применяются: принципы самоподобия для нестационарного трафика, который нарушается при возникновении аномальной активности в сети; методы машинного обучения для стационарного трафика.

При экспериментальной проверке разработанной модели для нахождения показателя Херста использовались *R/S* анализ и метод *DFA*. Для расчета показателя Херста в нестационарном трафике на малых выборках используется *R/S* анализ, а на зашумленных или больших объемах данных используется *DFA* анализ (рис. 2).

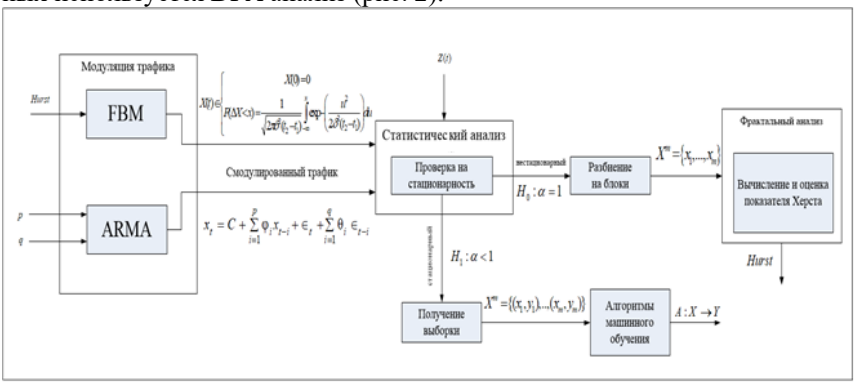

Рисунок 1 – Структура модели выявления аномалий в сетевом трафике СПД в условиях КА

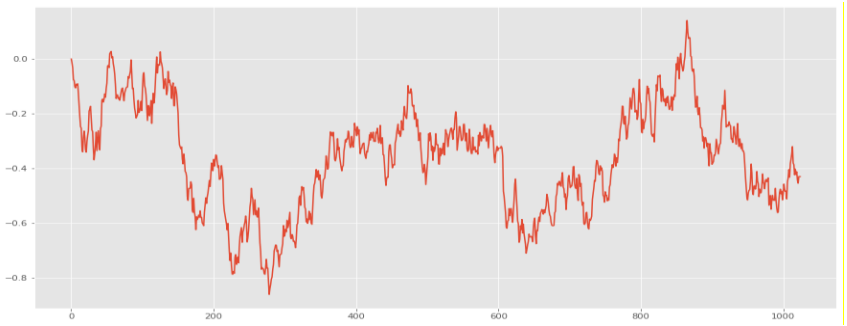

Рисунок 2 – Фрактальное броуновское движение при H = 0,5

Процесс, представленный на рис. 2, смоделирован с заданным показателем Херста (H=0,5). С помощью метода *DFA* произведен расчет показателя Херста (рис. 3) и произведено сравнение с эталонным значением. Для этого использовалась программа, написанная на *python*3, а для построения (отрисовки) графиков – графическая библиотека *Matplotlib*.

Затем выполнялся анализ самоподобия *H* смоделированного сигнала и найденное значение параметра *H* сравнивалось с эталонным.

С помощью разработанного киберполигона исследовалась эталонная выборка сетевого трафика СПД, которая включала в себя порядка 30 видов КА и 40 Гб легитимного сетевого трафика СПД, который перенаправлялся на *DPI – S ecurity Onion* и записывался в файлы с расширением *.pcap*.

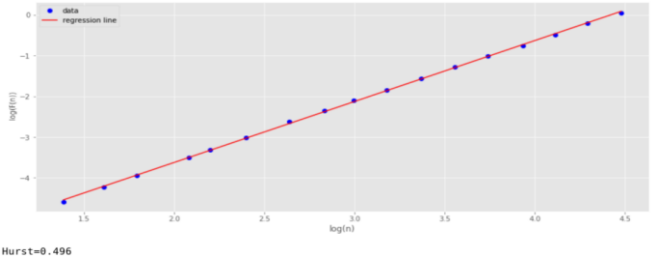

Рисунок 3 – Зависимость $F(n)$ от времени в логарифмической шкале (H = 0,496)

Из этого сетевого трафика СПД формировался сетевой набор данных с помощью анализаторов *Netsniff-ng* и *Bro*. КА проводились с помощью дистрибутива *Kali Linux* на заведомо уязвимые сервисы.

После сбора сетевого трафика СПД осуществлялась проверка его на стационарность. В случае нестационарности сетевого трафика – применялся фрактальный анализ. На рис. 4 представлена выборка из 1200 сетевых пакетов.

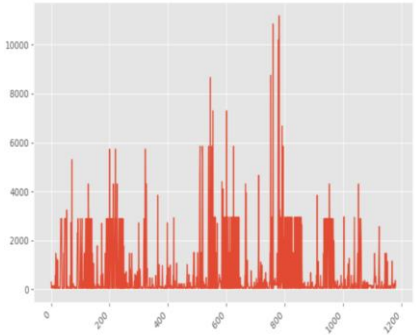

Рисунок 4 – Выборка из 1200 сетевых пакетов

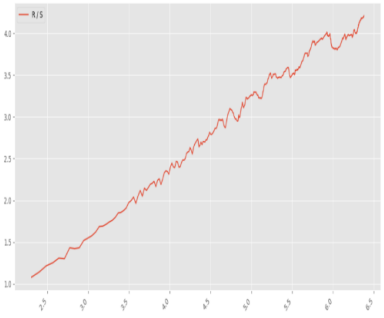

Рисунок 5 – *R/S*-анализ: оценка из выборки из 1200

С помощью *R/S* анализа в выборке из 1200 сетевых пакетов был найден показатель Херста (рис. 5). Его значение равно 0,74, что подтверждает наличие фрактальных свойств у сетевого трафика СПД. В работе

были проведены аналогичные исследования с сетевым набором данных объемом 40Гб, собранным за год.

Объем исследуемого сетевого трафика СПД с аномалиями равнялся 40Гб (рис. 6). С помощью *R/S* анализа был найден *H*=1,38, что указывает на наличие аномалий (рис. 7).

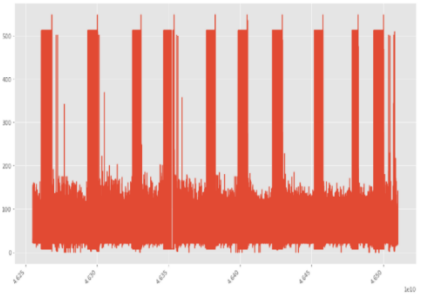

Рисунок 6 – Аномальный сетевой трафик СПД

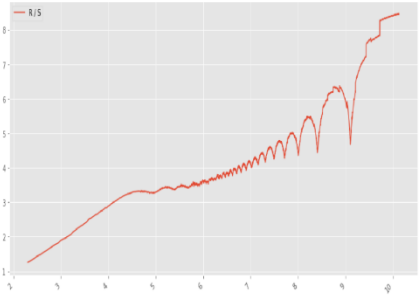

Рисунок 7 – *R/S*-анализ: оценка

Таким образом, разработана аналитическая модель выявления аномалий в сетевом трафике СПД в условиях КА, в которой в качестве основных методов оценки самоподобия используются *R/S* анализ и метод *DFA*. Проведено исследование с аномальным сетевым трафиком, содержащим *DoS* пакеты (рис. 6).

При тестировании фрактальных методов, позволяющих проводить исследования долговременных зависимостей в сетевом трафике СПД, было выявлено, что метод *DFA* является более эффективным, чем *R/S* анализ из-за его возможности обрабатывать зашумленные сигналы и большие объемы данных, что позволяет обнаруживать корреляции на больших расстояниях. Но в тоже время он уступает *R/S* в скорости. Так как реагирования на КА должно быть максимально быстрым, принято решение разбивать сетевой трафик СПД на интервалы и использовать *R/S* анализ.

В третьем разделе рассматривается методика раннего обнаружения аномалий в сетевом трафике СПД (рис. 8), позволяющая обнаруживать КА на раннем этапе их проявления с помощью методов машинного обучения для стационарного сетевого трафика СПД и фрактального анализа для нестационарного. Методика состоит из следующих этапов: сбор сетевого трафика в СПД; осуществляется его проверка на стационарность; подготовка исходных данных; фрактальный анализ для нестационарного сетевого трафика в СПД; машинное обучение для стационарного сетевого трафика в СПД. После сбора сетевого трафика осуществляется его проверка на стационарность. В случае нестационарности сетевого трафика СПД применяется фрактальный анализ.

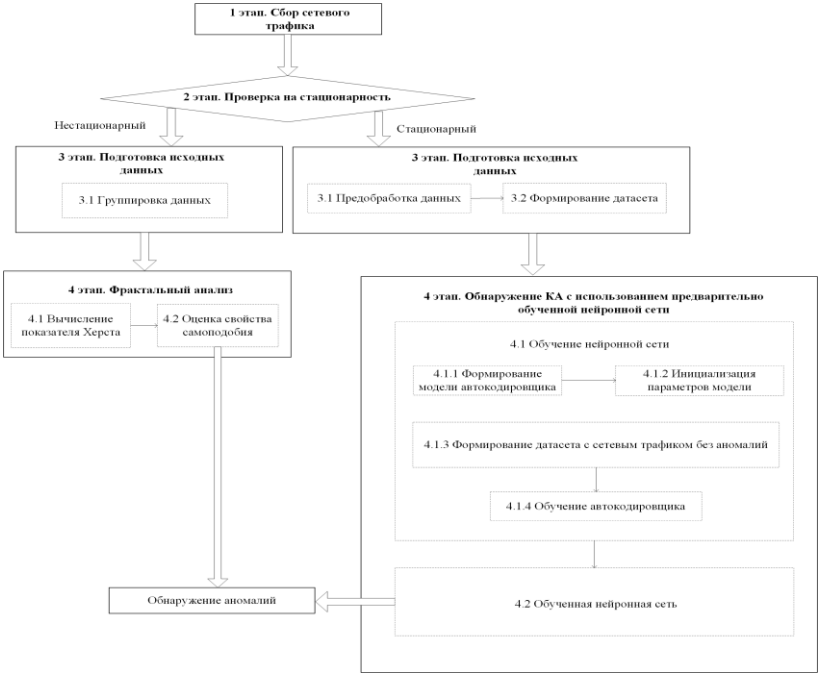

Рисунок 8 – Блок-схема методики раннего обнаружения аномалий в сетевом трафике СПД

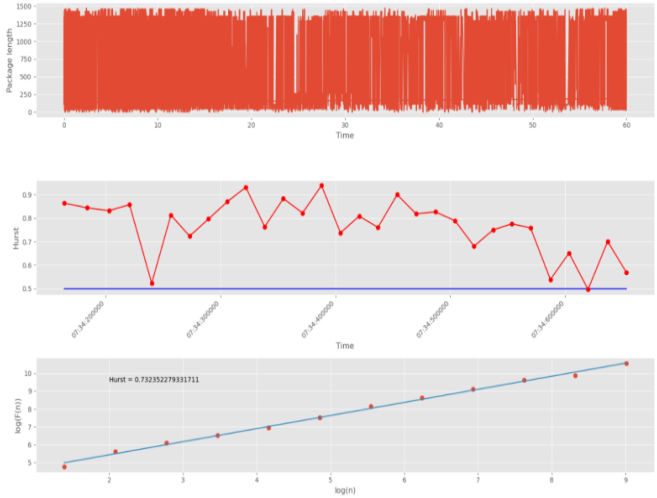

Рисунок 9 – Вычисление *H* методом фрактального анализа легитимного *UDP* трафика. Разбиение 10000 точек на 29 групп

Сетевой поток делится на группы, для каждой из которых рассчитывается показатель Херста. Сетевые пакеты помечаются аномальными в том случае, когда нарушается свойство самоподобия в исследуемой группе. Для проверки эффективности предложенного подхода, сперва алгоритм протестирован на легитимном сетевом трафике (рис. 9).

Синей прямой линией обозначен порог, соответствующий границе белого шума (H = 0,5). Точки на втором графике соответствуют номерам групп пакетов (всего 30 точек). Точки на третьем соответствуют *scales* (всего 12 точек).

Как видно из рис. 9, мера фрактальности для всех групп пакетов полностью лежит выше отметки 0,5. Это указывает на наличие самоподобных свойств у каждой группы. Кроме того, на третьем графике (логарифмической регрессии) отражен параметр Херста для всего сетевого набора данных, который подтверждает наличие фрактальных свойств и повторяющихся процессов. Далее проводилось тестирование аномального сетевого трафика СПД, полученного во время проведения *DoS* атаки и компьютерно-технической разведки.

Также в работе проведен анализ наиболее распространённых методов машинного обучения, направленных на обнаружение КА в стационарном сетевом трафике СПД, который показал, что несмотря на высокую точность обнаружения КА, современные классификаторы имеют существенный недостаток – большое количество ложных срабатываний. В работе также проведен анализ алгоритмов, основанных на математической статистике, который показал, что они отлично справляются с обнаружением аномальных всплесков.

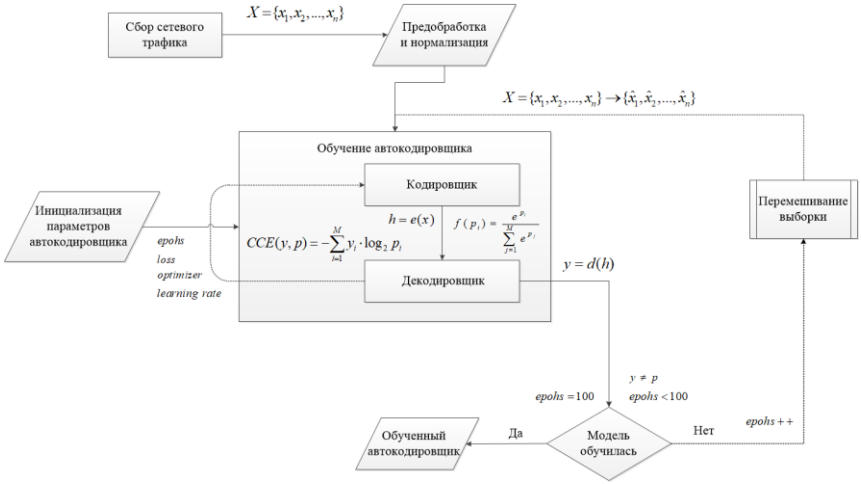

Рисунок 10 – Блок-схема методики обнаружения аномалий в стационарном сетевом трафике СПД

Однако не всегда всплески являются показателями аномалий. Именно поэтому для нахождения аномалий в стационарном сетевом трафике СПД в работе разработана соответствующая методика, в основу которой заложено использование нейронной сети – автокодировщика (рис 10).

Обученная нейронная сеть способна с минимальными потерями декодировать информацию схожего типа с той, на основе которой она обучалась. Если обученный автокодировщик встречает аномальный образец, то он воссоздает его с высокой степенью ошибки. Таким образом, ошибка восстановления между исходными и восстановленными данными будет выше для аномальных данных, чем для обычных (рис. 11).

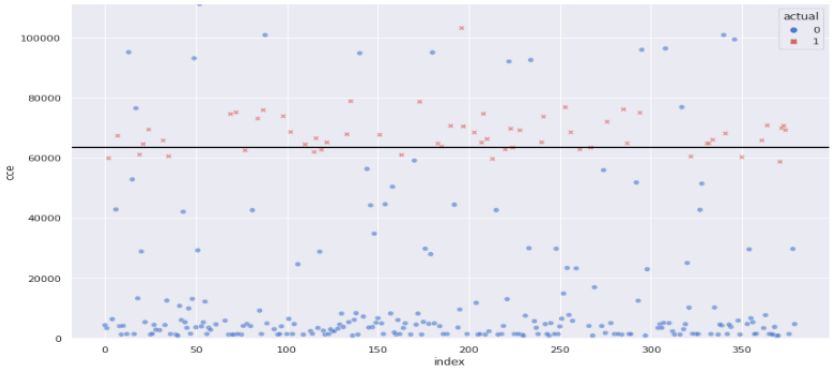

Рисунок 11 – Результат работы методики

Для оценки точности алгоритма построена *ROC* кривая и матрица ошибок первого и второго рода. Проведена оценка возможности раннего обнаружения КА в СПД с помощью классификаторов, методов математической статистики и автокодировщика, построенного на основе нейронной сети с долгой краткосрочной памятью.

Таким образом, проведена экспериментальная оценка методики раннего обнаружения аномалий в сетевом трафике, основанной на автокодировщике. Результаты экспериментальной проверки позволяют сделать вывод, что предложенная методика является достаточно корректной.

В четвертом разделе рассматривается методика классификации КА в сетевом трафике СПД (рис. 12), позволяющая выявлять КА с использованием гибридной нейронной сети, состоящей из классификатора и автокодировщика, обученного на основе данных работы функционирования СПД, учитывающего все отклонения от ее штатной работы. В процессе работы классификатор дополнительно обучается на скрытых латентных представлениях полученных автокодировщиком, т.е. в итоге получается генеративно-состязательная сеть, в которой нейронные сети учатся друг у друга.

Сбор данных, предназначенных для обучения нейронной сети, осуществлялся с помощью промежуточного программного слоя (*middleware*).

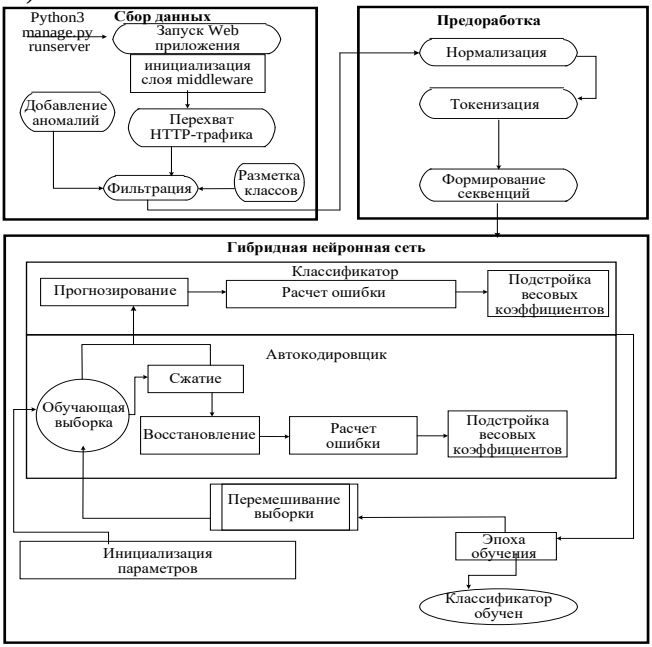

Рисунок 12 – Блок-схема методики классификации КА в сетевом трафике СПД

Такой подход позволяет обрабатывать запросы из браузера прежде, чем они достигнут представления *Django* (сервера), а также ответы от представлений до того, как они возвращаются в браузер. Перехваченные запросы записываются в лог-файл (журнализируются).

В набор данных, сформированный после сбора *HTTP* трафика, добавляются аномальные запросы для задания мультиклассовости КА. К такого рода запросам относятся: *SQL Injection*, *Cross-Site Scripting*, *Cross-Site Request Forgery*, *XML External Entity Injection*, *CRLF Injection* и *HTTP Response Splitting*. Каждый тип аномального запроса помечается в наборе данных как отдельный класс.

Следующим этапом является предобработка сформированного набора данных, главной задачей которой является нормализация данных. После нормализации данные токенизируются. Для этой цели сперва осуществляется замена символов, встречающихся в наборе данных, на числовой эквивалент, который не имеет самостоятельного значения для внешнего или внутреннего использования. Затем слова переводятся в последова-

тельность секвенций. При этом важным обстоятельством является то, что все секвенции должны быть одной длины. Если запрос меньше длины секвенции, то оставшиеся символы заполняются нулями.

После этого инициализируются гиперпараметры гибридной нейронной сети и происходит ее обучение на сформированных секвенциях. Классификатор гибридной нейронной сети ищет закономерности в данных, привязывая их к размеченным классам на этапе обучения.

Данные, которые классификатор не смог отнести к какому-либо классу, в том числе и к классу легитимных данных, помечаются как аномальные. Они соответствуют атакам «нулевого дня». Для классификации аномалий в стационарной сети предлагается использовать гибридную нейронную сеть (рис. 13).

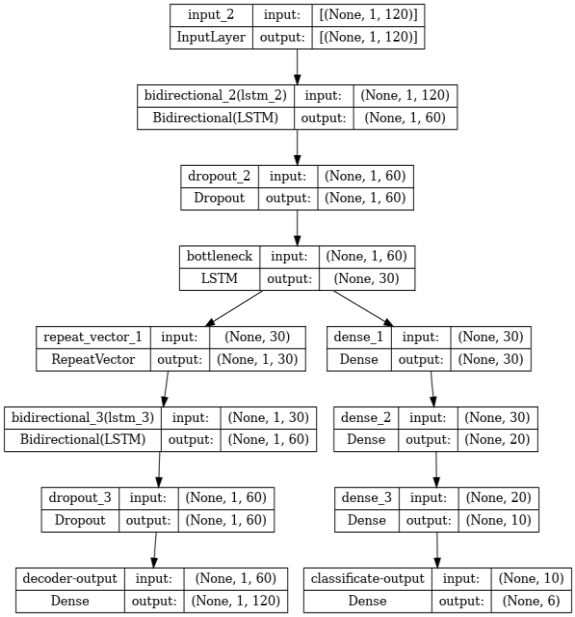

Рисунок 13 – Модель гибридной нейронной сети

У нейронной сети имеется два выхода. Левая ветвь на рис. 20 соответствует автокодировщику, а правая – классификатору. Гибридная сеть имеет различные по своему назначению слои. Входной слой гибридной нейронной сети имеет 120 нейронов, применяющиеся как для автокодировщика, так и для классификатора.

В качестве слоев автокодировщика используются ячейки с долгой краткосрочной памятью, т.е. *long short-term memory* (LSTM), которая может удалять информацию из состояния ячейки. Этот процесс регулируется фильтрами. Они позволяют пропускать информацию на основании некоторых условий.

В процессе обучения гибридной нейронной сети на входной слой поступают вектора (падасеквенции, секвенции или эмбеддинги) в зависимости от реализации. Эксперименты показали, что наилучший результат дают эмбеддинги, но они сильно нагружают ЭВМ, поэтому для быстрых

вычислений в исследованиях применялись падасеквенции длиной 120 символов.

В середине гибридной нейронной сети число нейронов уменьшается до 30. Это приводит к потере информации, так как из 120 нейронов не вся информация попадает на 30 нейронов. Поэтому нейронная сеть учится отбрасывать лишнюю (избыточную) информацию (шум), сохраняя, по ее мнению, самую важную.

У классификатора (правая ветвь) на последнем слое имеется шесть нейронов. Они соответствуют шести размеченным классам, упомянутым выше. В случае, если классификатор не может отнести данные ни к одному из классов с вероятностью больше 0,6, то такой запрос помечается подозрительным (считается атакой нулевого дня).

У автокодировщика (левая ветвь) на последнем слое 120 нейронов (столько же, сколько и на входном). Он пытается воссоздать информацию из 30 нейронов к первоначальному виду. Каждый раз в процессе обучения у автокодировщика получается делать это все лучше и лучше.

Таким образом, уменьшается ошибка, а в середине слоя сохраняется только самая важная информация, из которой можно восстановить информацию в исходном виде – «скрытые латентные представления». Такие представления позволили классификатору находить дополнительные закономерности в данных, что существенно уменьшает ложные срабатывания и повышает вероятность обнаружения атак нулевого дня.

После обучения нейронной сети, была проведена экспериментальная оценка точности и полноты с применением ROC-кривой на данных, не участвующих в обучении модели. Визуальный анализ подтвердил высокую точность предложенного подхода, с минимальным числом ложных срабатываний.

Для более качественной оценки выделены главные метрики классификации (рис. 14.

	precision	recall	f1-score	support
xxe	0.83	0.95	0.88	1292
no anomaly	0.99	1.00	0.99	9266
server side template injection	1.00	0.99	0.99	1265
sql injection	0.96	0.93	0.95	1271
traversal	0.94	0.80	0.86	1263
encoding traversal	0.99	0.99	0.99	1305
accuracy			0.97	15662
macro avg	0.95	0.94	0.94	15662
weighted avg	0.97	0.97	0.97	15662

Рисунок 14 – Главные метрики классификаций

Из рисунка 14 видно, что алгоритм не только справляется с классификацией КА, но также с высокой долей вероятности обнаруживает легитимные запросы «no anomaly», что само по себе является минимизацией

ложных срабатываний (ошибок первого рода), связанных с отнесением легитимных запросов к КА и наоборот. Таким образом классификация КА подтверждает отсутствие ложных срабатываний при мультиклассовой классификации аномалий.

Также представлена архитектура и программный прототип компонентов системы раннего обнаружения и классификации КА в сетевом трафике СПД, которые отличаются от известных тем, что ориентированы на мгновенное обнаружение как известных, так и неизвестных КА, их классификацию и выбор доступных контрмер с минимальным числом ложных срабатываний (рис. 15).

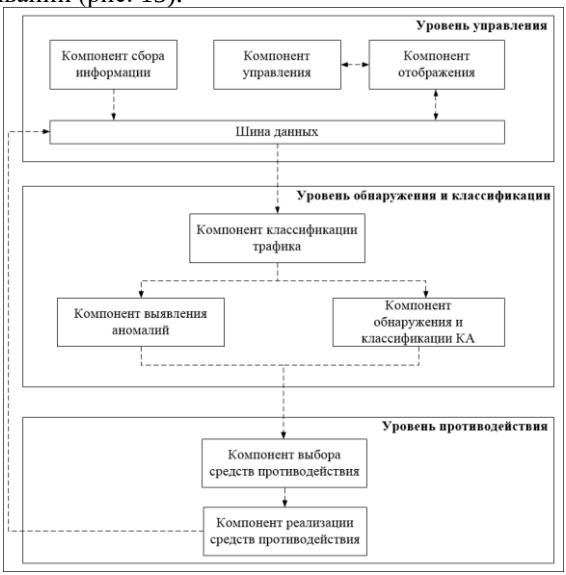

Рисунок 15 – Архитектура системы раннего обнаружения и классификации КА в сетевом трафике СПД

Архитектура включает три уровня: (1) уровень управления (компонент сбора информации, компонент управления, компонент отображения); (2) уровень обнаружения и классификации (компонент классификации сетевого трафика СПД, компонент выявления аномалий, а также компонент обнаружения и классификации КА); (3) уровень противодействия (компонент выбора средств противодействия и компонент реализации средств противодействия). Она позволяет формировать наборы исходных данных для исследований и обучения гибридной нейронной сети, выявлять известные и неизвестные КА в реальном масштабе времени, вырабатывать решения по реализации средств защиты, уведомлять администратора безопасности о реализации КА и принятых мерах.

Полученная архитектура является разработкой в области раннего обнаружения и противодействия КА в СПД, а также исследований и разработок решений для систем поддержки принятия решения. Элементы архитектуры реализованы в качестве программного компонента обнаружения аномалий сетевого трафика СПД на основе принципов фрактального анализа данных и методах машинного обучения.

Результаты эксперимента показали, что разработанные в диссертационной работе методики обладают достаточно высокой эффективностью при обнаружении как известных, так и неизвестных КА. Вероятность обнаружения известных КА составляет 0,96, а неизвестных атак – 0,8.

Результаты сравнительной оценки полученных результатов с сигнатурными и статистическими методами, а также методами машинного обучения, продемонстрировали высокую скорость обнаружения и точность обнаружения известных и неизвестных типов КА (табл. 1).

Табл. 1. Сравнительный анализ методов обнаружения КА.

Наименование методов	Скорость обнаружения (сек)	Точность обнаружения КА		Ложные срабатывания		Тип трафика	
		Известных	Неизвестных	Бинарная классификация	Мультиклассовая классификация	Стационарный	Нестационарный
Сигнатурные методы	5	0,99	0,5	-	-	+	-
Статистические методы	30	0,92	0,6	-	-	+	-
Методы маш. Обуч.	28	0,72-0,97	0,6	0,39-0,56	0,06-0,92	+	-
Разраб. методики	5	0,96	0,8	0,05	0-0,20	+	+

В качестве основных учитываемых параметров сравниваемых методов рассматривались скорость и точность обнаружения КА, как известных, так и неизвестных, возможность работы со стационарным и нестационарным трафиком, а также то, насколько часто происходят ложные срабатывания.

Из таблицы 1 видно, что разработанные методики по скорости не уступают самым лучшим известным подходам (сигнатурным методам), по точности превышают все известные методы, а также имеют более широкую область применения (охватывают нестационарный трафик), т.е. можно сделать вывод, что эффективность обнаружения КА повысилась.

Сигнатурные методы и предлагаемые методики являются самыми быстрыми с точки зрения скорости обнаружения. Кроме того, поскольку сигнатурные методы используют предопределенные правила, они обладают высочайшей точностью обнаружения известных атак. Однако их точность в обнаружении неизвестных атак очень низка. Значение 0,5 указывает на то, что эта точность соответствует закону равной вероятности.

Статистические методы проигрывают сигнатурным методам с точки зрения скорости обнаружения и точности, поскольку они используют накопленную статистику. Однако иногда они способны обнаруживать неизвестные атаки.

Методы машинного обучения довольно разнообразны и хорошо развиты. Их эффективность зависит от используемых ими моделей классификации и кластеризации. Методы машинного обучения проигрывают сигнатурным методам в скорости обнаружения. Однако они обладают более высокой точностью обнаружения неизвестных атак.

Разработанные методики имеют скорость обнаружения, подобную сигнатурным методам, а точность соответствует значениям, подобным методам машинного обучения. В то же время они сохраняют свою эффективность при работе с различными типами сетевого трафика СПД.

Таким образом, полученные в диссертации результаты позволяют утверждать о достижении более высокой эффективности разработанной методики по сравнению с известными, что доказывает реализацию итоговой цели исследования – повышение эффективности выявления аномалий и классификации КА в сетевом трафике СПД.

В заключении перечислены полученные теоретические и практические результаты, раскрыта степень их новизны и значение для теории и практики, предложены перспективные направления дальнейших исследований, ориентированных на обеспечение защиты СПД.

ЗАКЛЮЧЕНИЕ

В диссертационной работе решена научная задача, заключающаяся в разработке модели и методик выявления аномалий и классификации КА в сетевом трафике СПД на основе применения фрактального анализа и методов машинного обучения.

Решенная задача имеет важное значение для совершенствования моделей, методик и средств раннего выявления аномалий в сетевом трафике СПД, находящихся под воздействием как известных, так и неизвестных КА, а также прогнозирования факта их воздействия.

Основные научные результаты, составляющие **итоги** выполненного исследования:

1) Проанализированы существующие модели воздействия КА на СПД.

2) Проанализированы существующие алгоритмы выявления КА, существующих систем мониторинга и методик противодействия КА в СПД.

3) Разработана аналитическая модель выявления аномалий в сетевом трафике СПД в условиях КА.

4) Разработана методика раннего обнаружения аномалий в сетевом трафике СПД.

5) Разработана методика классификации КА в сетевом трафике СПД.

6) Разработаны архитектуры и программные прототипы компонентов системы раннего обнаружения и классификации КА в сетевом трафике СПД.

7) Проведена экспериментальная и теоретическая оценка эффективности предложенных модели, методик и архитектуры, а также сравнение с существующими методиками.

Научная новизна результатов исследования заключается в том, что:

разработанная аналитическая модель выявления аномалий в сетевом трафике СПД в условиях КА, в отличие от известных, описывает не только стационарный, но и нестационарный сетевой трафик и обосновывает выбор фрактального метода выявления аномалий в интересах дальнейшей классификации КА в зависимости от типа трафика;

разработанная методика раннего обнаружения аномалий в сетевом трафике СПД, в отличие от известных, способна проводить прогнозирование и обнаружение не только известных, но и неизвестных КА на раннем этапе их проявления благодаря совместному применению методов фрактального анализа и искусственной нейронной сети LSTM-типа (рекуррентные нейронные сети с долгой краткосрочной памятью), в результате чего существенно сокращается время выявления аномалий и уменьшается количество ложных срабатываний;

разработанная методика классификации КА в сетевом трафике СПД отличается от известных тем, что в ней обнаружение КА производится с использованием генеративно-состязательной сети, которая производит дообучение классификатора на скрытых латентных представлениях, полученных автокодировщиком, исходя из анализа данных о функционировании СПД;

предложенная архитектура и программные прототипы компонентов системы раннего обнаружения и классификации КА в СПД отличаются от известных тем, что они ориентированы на раннее обнаружение как известных, так и неизвестных КА с минимальным количеством ложных срабатываний за счет реализации методов фрактального анализа и искусственной нейронной сети LSTM-типа.

Все выносимые на защиту результаты являются новыми и получены соискателем самостоятельно. Совокупное применение разработанных модели и методик, по сравнению со многими другими подходами, позволит увеличить скорость обнаружения КА в 14 раз путем выявления аномалий в трафике любого вида. Также предложенная система продемонстрировала достаточно высокую вероятность обнаружения КА, достигнув значения 0.96 для известных атак и 0.8 для ранее неизвестных атак. Даны рекомендации по использованию результатов исследования для повышения защищённости СПД от КА. Разработанные методики и модель направлены на сокращение времени выявления аномалий и классификации компьютерных

атак в сетях передачи данных, и могут быть применены на предприятиях промышленности при выполнении научных исследований.

Рекомендации. Разработанные алгоритмы могут быть использованы в существующих системах глубокого анализа сетевого трафика (IDS и IPS), системах обнаружения атак, поскольку они представляют собой инструмент моментального выявления аномалий в сетевом трафике

Перспективы дальнейшей разработки темы. В качестве перспектив дальнейшей разработки темы можно указать исследования, связанные с интеграцией предлагаемой системы с другими известными системами защиты и имеющимися в арсенале систем компьютерной безопасности методами детектирования атак, а также апробацию разработанного программного компонента обнаружения аномалий в сетевом трафике на основе принципов фрактального анализа данных на принципиально других типах СПД.

Соответствие специальности. Полученные результаты соответствуют специальности 2.3.6. «Методы и системы защиты информации, информационная безопасность».

СПИСОК РАБОТ ПО ТЕМЕ ДИССЕРТАЦИИ

1. Kotenko, I., Saenko, I., Lauta, O., Kribel, A. A proactive protection of smart power grids against cyberattacks on service data transfer protocols by computational intelligence methods // Sensors 2022, 22, 7506.
2. Kotenko, I., Saenko, I., Lauta, O., Kribel, A. An approach to detecting cyber attacks against smart power grids based on the analysis of network traffic self-similarity // Energies, 2020, 13(19), 5031.
3. Kribel, A., Saenko, I., Kotenko, I. Detection of Anomalies in the Traffic of Information and Telecommunication Networks Based on the Assessment of its Self-Similarity // Proceedings - 2020 International Russian Automation Conference, RusAutoCon 2020, 2020, стр. 713–718, 9208147.
4. Kotenko, I., Saenko, I., Kribel, A., Lauta, O. A technique for early detection of cyberattacks using the traffic self-similarity property and a statistical approach // Proceedings - 29th Euromicro International Conference on Parallel, Distributed and Network-Based Processing, PDP 2021, 2021, стр. 281–284, 9407132.
5. Kotenko, I., Saenko, I., Lauta, O., Kribel, A. Ensuring the survivability of embedded computer networks based on early detection of cyber attacks by integrating fractal analysis and statistical methods // Microprocessors and Microsystemsthis link is disabled, 2022, 90, 104459.
6. Kotenko, I., Saenko, I., Lauta, O., Kribel, A. Anomaly and cyber attack detection technique based on the integration of fractal analysis and machine learning methods // Informatics and Automationthis link is disabled, 2022, 21(6), pp. 1328–1358.

7. Котенко И.В., Саенко И.Б., Лаута О.С., Крибель А.М. Метод раннего обнаружения кибератак на основе интеграции фрактального анализа и статистических методов // Первая миля. 2021. № 6 (98). С. 64-71.

8. Крибель А.М., Лаута О.С., Филин А.В., Фень А.С. Метод обнаружения аномалий в сетевом компьютерном трафике на основе нейронной сети с использованием LSTM // Электросвязь. 2021. № 12. С. 43-48.

9. Саенко И.Б., Лаута О.С., Карпов М.А., Крибель А.М. Модель угроз ресурсам ИТКС как ключевому активу критически важного объекта инфраструктуры // Электросвязь. 2021. № 1. С. 36-44;

10. Котенко И.В., Крибель А.М., Лаута О.С., Саенко И.Б. Анализ процесса самоподобия сетевого трафика как подход к обнаружению кибератак на компьютерные сети // Электросвязь. 2020. № 12. С. 54-59.;

11. Крибель А.М. Методика обнаружения коллизий сетевого трафика // Известия Тульского государственного университета. Технические науки. 2021. № 12. С. 182-190.

12. Панков А.В., Крибель А.М., Лаута О.С., Васильев Н.А. Метод по совершенствованию информационно- аналитической работы на основе комплексирования результатов распознавания состояний объектов контроля с использованием методов машинного обучения // Наукоемкие технологии в космических исследованиях Земли. 2022. Т. 14. № 2. С. 27-35.

13. Перов Р.А., Лаута О.С., Крибель А.М., Федулов Ю.М. Комплексная методика обнаружения кибератак на основе интеграции фрактального анализа и статистических методов // Наукоемкие технологии в космических исследованиях Земли. 2022. Т. 14. № 2. С. 44-51.

14. Крибель А.М., Перов Р.А., Лаута О.С., Сычужников В.Б. Методика обнаружения компьютерных атак с помощью фрактального анализа и методов машинного обучения // Известия Тульского государственного университета. Технические науки. 2022. № 5. С. 166-178.

15. Крибель А.М., Перов Р.А., Лаута О.С., Скоробогатов С.Ю. Модель выявления аномалий в сетевом трафике сети передачи данных в условиях компьютерных атак // Известия Тульского государственного университета. Технические науки. 2022. № 5. С. 228-239.

16. Перов Р.А., Лаута О.С., Крибель А.М., Федулов Ю.В. Метод выявления аномалий в сетевом трафике // Наукоемкие технологии в космических исследованиях Земли. 2022. Т. 14. № 3. С. 25-31.

17. Котенко И. В., Саенко И. Б., Лаута О. С., Крибель А.М. Программный компонент обнаружения аномалий сетевого трафика на основе принципов фрактального анализа данных // Свидетельство о регистрации программы для ЭВМ 2021680188, 07.12.2021.

Автореферат диссертации

КРИБЕЛЬ
Александр Михайлович

ВЫЯВЛЕНИЕ АНОМАЛИЙ И КЛАССИФИКАЦИЯ КОМПЬЮТЕРНЫХ АТАК В СЕТИ ПЕРЕДАЧИ ДАННЫХ НА ОСНОВЕ ПРИМЕНЕНИЯ ФРАКТАЛЬНОГО АНАЛИЗА И МЕТОДОВ МАШИННОГО ОБУЧЕНИЯ

Текст автореферата размещен на сайтах:
Высшей аттестационной комиссии при Министерстве науки и высшего образования Российской Федерации
https://vak.minobrnauki.gov.ru/

Федерального государственного бюджетного учреждения науки «Санкт-Петербургский Федеральный исследовательский центр Российской академии наук»
http://www.spiiras.nw.ru/dissovet/

Подписано в печать "___" _____ 2023 г.
Формат 60x84 1/16. Бумага офсетная. Печать офсетная.
Усл.печ.л. 1,0. Тираж 100 экз.
Заказ № ___